

Inverse Problems

PAPER • OPEN ACCESS

An alternating iterative minimisation algorithm for the double-regularised total least square functional

To cite this article: Ismael Rodrigo Bleyer and Ronny Ramlau 2015 *Inverse Problems* **31** 075004

View the [article online](#) for updates and enhancements.

Related content

- [A double regularization approach for inverse problems with noisy data and inexact operator](#)
Ismael Rodrigo Bleyer and Ronny Ramlau
- [Sparsity regularization for parameter identification problems](#)
Bangti Jin and Peter Maass
- [A discrepancy-based parameter adaptation and stopping rule for minimization algorithms aiming at Tikhonov-type regularization](#)
Kristian Bredies and Mariya Zhariy

Recent citations

- [Automatic alignment for three-dimensional tomographic reconstruction](#)
Tristan van Leeuwen *et al*
- [Alternating minimisation for glottal inverse filtering](#)
Ismael Rodrigo Bleyer *et al*

An alternating iterative minimisation algorithm for the double-regularised total least square functional

Ismael Rodrigo Bleyer¹ and Ronny Ramlau²

¹ Department of Mathematics and Statistics, University of Helsinki, Gustaf Hällströmin Katu 2b, FI-00014 Helsinki, Finland

² Industrial Mathematics Institute, Johannes Kepler University Linz, Altenbergerstraße 69, A-4040 Linz, Austria

E-mail: ismael.bleyer@helsinki.fi and ronny.ramlau@jku.at

Received 15 August 2014, revised 23 April 2015

Accepted for publication 24 April 2015

Published 19 May 2015



CrossMark

Abstract

The *total least squares* (TLS) method is a successful approach for linear problems if both the right-hand side and the operator are contaminated by some noise. For ill-posed problems, a regularisation strategy has to be considered to stabilise the computed solution. Recently a double regularised TLS method was proposed within an infinite dimensional setup and it reconstructs both function and operator, reflected on the bilinear forms. Our main focuses are on the design and the implementation of an algorithm with particular emphasis on alternating minimisation strategy, for solving not only the double regularised TLS problem, but a vast class of optimisation problems: on the minimisation of a bilinear functional of two variables.

Keywords: ill-posed problems, noisy operator, noisy right-hand side, regularised total least squares, alternating minimisation, wavelets, subderivatives

(Some figures may appear in colour only in the online journal)



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

In [2], the authors described a new two-parameter regularisation scheme for solving an ill-posed operator equation. The task consists of the inversion of a linear operator $A_0: \mathcal{V} \rightarrow \mathcal{H}$ defined between Hilbert spaces

$$A_0 f = g_0. \quad (1)$$

In contrast to standard inverse problems, where the task is to solve (1) from given noisy data, a more realistic setup is considered where both data and operator are not known exactly. For the reconstruction, a cost functional with two penalisation terms based on the TLS (total least squares) technique is used.

This approach presented in [2] focuses on linear operators that can be characterised by a function, as it is, e.g. the case for linear integral operators, where the kernel function determines the behaviour of the operator. Moreover, it is assumed that the noise in the operator is due to an incorrect characterising function. A penalty term is not only used to stabilise the reconstruction of the unknown solution, as it is the case in [10–12], but also to stabilise the unknown operator. As a drawback, the regularisation scheme becomes nonlinear even for linear equations. However, the potential advantage is that not only the unknown solution is reconstructed, but also a suitable characterising function and thus the governing operator describing the underlying data. Additionally, convergence rates for the reconstruction of both solution and operator have been derived.

The *double regularised total least squares* (dbl-RTLS) approach allow us to treat the problem in the framework of Tikhonov regularisation rather than as a constraint minimisation problem. More precisely, the regularised solution is obtained by minimising a nonlinear, nonconvex and possibly non-differentiable functional over two variables, which is computationally not always straightforward. Thus the goal of this paper is the development of an efficient and convergent numerical scheme for the minimisation of the Tikhonov-type functional for the dbl-RTLS approach.

The rest of paper is organised as follows: in section 2 we formulate the underlying problem and give a short summary of the dbl-RTLS method. Section 3 is dedicated to the development of an algorithm based on an alternating minimisation strategy, as well as its convergence properties. In section 4, numerical results for the proposed algorithm are provided and the efficiency of the method is discussed. For the convenience of the reader in appendix we display important concepts and fundamental results used throughout this article.

2. Problem formulation and the dbl-RTLS method

As mentioned above, we aim at the inversion of the linear operator equation (1) from noisy data g_δ and an incorrect operator A_ϵ . Additionally we assume that the operators $A_0, A_\epsilon: \mathcal{V} \rightarrow \mathcal{H}$, where \mathcal{V} and \mathcal{H} are Hilbert spaces, can be characterised by functions $k_0, k_\epsilon \in \mathcal{U}$, \mathcal{U} also a Hilbert space. To be more specific, we consider operators

$$\begin{aligned} A_k: \mathcal{V} &\longrightarrow \mathcal{H} \\ v &\longmapsto B(k, v), \end{aligned}$$

i.e. $A_k v := B(k, v)$, where B is a *bilinear* operator

$$B: \mathcal{U} \times \mathcal{V} \rightarrow \mathcal{H}$$

fulfilling, for some $C > 0$, the inequality

$$\|B(k, f)\|_{\mathcal{H}} \leq C \|k\|_{\mathcal{V}} \|f\|_{\mathcal{V}}. \quad (2)$$

From (2) follows immediately

$$\|B(k, \cdot)\|_{\mathcal{V} \rightarrow \mathcal{H}} \leq C \|k\|_{\mathcal{V}}. \quad (3)$$

Associated to the bilinear operator B , we also define the linear operator

$$\begin{aligned} C_f: \mathcal{U} &\longrightarrow \mathcal{H} \\ u &\longmapsto B(u, f), \end{aligned}$$

i.e. $C_f u := B(u, f)$.

From now on, let us identify A_0 with A_{k_0} and A_ϵ with A_{k_ϵ} . From (3) we deduce immediately

$$\|A_0 - A_\epsilon\| \leq C \|k_0 - k_\epsilon\|, \quad (4)$$

i.e. the operator error norm is controlled by the error norm of the characterising functions. Now we can formulate our problem as follows:

$$\text{Solve } A_0 f = g_0 \quad (5a)$$

$$\text{from noisy data } g_\delta \text{ with } \|g_0 - g_\delta\| \leq \delta \quad (5b)$$

$$\text{and noisy function } k_\epsilon \text{ with } \|k_0 - k_\epsilon\| \leq \epsilon. \quad (5c)$$

Please note that the problem with explicitly known k_0 (or the operator A_0) is often ill-posed and needs regularisation for a stable inversion. Therefore we will also propose a regularising scheme for the problem (5a)–(5c).

Due to our assumptions on the structure of the operator A_0 , the inverse problem of identifying the function f^{true} from noisy measurements g_δ and an inexact operator A_ϵ can now be rewritten as the task of solving the inverse problem find f s.t.

$$B(k_0, f) = g_0 \quad (6)$$

from noisy measurements (k_ϵ, g_δ) fulfilling

$$\|g_0 - g_\delta\|_{\mathcal{H}} \leq \delta, \quad (7a)$$

and

$$\|k_0 - k_\epsilon\|_{\mathcal{V}} \leq \epsilon. \quad (7b)$$

In most applications, the ‘inversion’ of B will be ill-posed (e.g. if B is defined via a Fredholm integral operator), and a regularisation strategy is needed for a stable solution of the problem (6).

For the solution of (6) from given data (k_ϵ, g_δ) fulfilling (7), we use the dbl-RTLS method proposed in [2], where the approximations to the solutions are computed as

$$\arg \min_{(k, f)} J_{\alpha, \beta}^{\delta, \epsilon}(k, f) := \frac{1}{2} T^{\delta, \epsilon}(k, f) + R_{\alpha, \beta}(k, f), \quad (8a)$$

where

$$T^{\delta, \epsilon}(k, f) = \|B(k, f) - g_\delta\|^2 + \gamma \|k - k_\epsilon\|^2 \quad (8b)$$

and

$$R_{\alpha,\beta}(k, f) = \frac{\alpha}{2} \|Lf\|^2 + \beta\mathcal{R}(k). \quad (8c)$$

Here, α and β are the regularisation parameters which have to be chosen properly, γ is a scaling parameter (arbitrary but fixed), L is a bounded linear and continuously invertible operator and $\mathcal{R}: X \subset \mathcal{U} \rightarrow [0, +\infty]$ is a proper, convex and weakly lower semi-continuous functional. The functional $J_{\alpha,\beta}^{\delta,\epsilon}$ is composed as the sum of two terms: one which measures the discrepancy of data and operator, and one which promotes stability. The functional $T^{\delta,\epsilon}$ is a *data-fidelity* term based on the TLS technique, whereas the functional $R_{\alpha,\beta}$ acts as a *penalty* term which stabilises the inversion with respect to the pair (k, f) . As a consequence, we have two regularisation parameters, which also occurs in *double regularisation*, see, e.g. [17].

The domain of the functional $J_{\alpha,\beta}^{\delta,\epsilon}: (\mathcal{U} \cap X) \times \mathcal{V} \rightarrow \mathbb{R}$ can be extended over $\mathcal{U} \times \mathcal{V}$ by setting $\mathcal{R}(k) = +\infty$ whenever $k \in \mathcal{U} \setminus X$. Then \mathcal{R} is proper, convex and weak lower semi-continuous functional in \mathcal{U} .

It has been shown that the sequence of the pair of solutions (k^n, f^n) of (8) converges to a *minimum-norm solution* when $(\delta, \epsilon) \rightarrow (0, 0)$, i.e. it is a regularisation method (see [2, theorem 4.5]). However, the task of finding minimisers of (8) has not been addressed properly, which will be done in the following sections.

3. An algorithm for the minimisation of the dbl-RTLS functional

In this section, we will formulate the first-order necessary condition for critical points of the functional $J_{\alpha,\beta}^{\delta,\epsilon}$, which requires in particular the derivative of the bilinear operator B . The core of this section is to design an algorithm to minimise $J_{\alpha,\beta}^{\delta,\epsilon}$, which is not a trivial task, as the functional is most likely nonconvex and nonlinear.

3.1. Optimality condition

It is well known that the study of local behaviour of nonsmooth functions can be achieved by the concept of *subdifferentiability* which replaces the classical derivative at non-differentiable points.

The first-order necessary condition based on subdifferentiability is stated as the following: if (\bar{k}, \bar{f}) minimises the functional $J_{\alpha,\beta}^{\delta,\epsilon}$ then

$$(0, 0) \in \partial J_{\alpha,\beta}^{\delta,\epsilon}(\bar{k}, \bar{f}). \quad (9)$$

We denote the set of all subderivatives of the functional $J_{\alpha,\beta}^{\delta,\epsilon}$ at (k, f) by $\partial J_{\alpha,\beta}^{\delta,\epsilon}(k, f)$ and we name it the *subdifferential* of $J_{\alpha,\beta}^{\delta,\epsilon}$ at (k, f) . For a quick revision on subdifferentiability we refer to the appendix.

The first result gives us the derivative of a bilinear operator B .

Lemma 3.1. *Let B be a bilinear operator and assume that (2) holds. Then the Fréchet derivative of B at $(k, f) \in \mathcal{U} \times \mathcal{V}$ is given by*

$$\begin{aligned} B'(k, f)(u, v) &= B(k, v) + B(u, f) \\ &= A_k v + C_f u. \end{aligned}$$

Moreover, the derivative is Lipschitz continuous with constant $\sqrt{2}C$.

Proof. We have to show

$$B(k + u, f + v) = B(k, f) + B'(k, f)(u, v) + o(\|(u, v)\|).$$

Since B is bilinear, we have

$$B(k + u, f + v) - B(k, f) = B(k, v) + B(u, f) + B(u, v),$$

and we observe $\|B(u, v)\| = o(\|(u, v)\|)$: As B fulfills (2), we have

$$\frac{\|B(u, v)\|}{\|(u, v)\|} \leq \frac{C \|u\| \|v\|}{(\|u\|^2 + \|v\|^2)^{1/2}} \leq \frac{C}{\sqrt{2}} (\|u\| \|v\|)^{1/2},$$

which converges to zero as $(u, v) \rightarrow 0$.

We further observe

$$\begin{aligned} B'(k, f)(u, v) - B'(\tilde{k}, \tilde{f})(u, v) &= B(k, v) + B(u, f) - (B(\tilde{k}, v) + B(u, \tilde{f})) \\ &= B(u, f - \tilde{f}) + B(k - \tilde{k}, v) \end{aligned}$$

which implies

$$\begin{aligned} \|B'(k, f)(u, v) - B'(\tilde{k}, \tilde{f})(u, v)\| &= \|B(u, f - \tilde{f}) + B(k - \tilde{k}, v)\| \\ &\leq \|B(u, f - \tilde{f})\| + \|B(k - \tilde{k}, v)\| \\ &\leq C \|u\| \|f - \tilde{f}\| + C \|k - \tilde{k}\| \|v\| \end{aligned}$$

Using the inequality $(a + b)^2 \leq 2(a^2 + b^2)$ we get

$$\begin{aligned} \|B'(k, f)(u, v) - B'(\tilde{k}, \tilde{f})(u, v)\|^2 &\leq 2C^2 (\|u\|^2 \|f - \tilde{f}\|^2 + \|k - \tilde{k}\|^2 \|v\|^2) \\ &\leq 2C^2 (\|u\|^2 + \|v\|^2) (\|k - \tilde{k}\|^2 + \|f - \tilde{f}\|^2) \\ &= 2C^2 \|(u, v)\|^2 \|(k - \tilde{k}, f - \tilde{f})\|^2 \end{aligned}$$

and thus

$$\begin{aligned} \|B'(k, f) - B'(\tilde{k}, \tilde{f})\| &= \sup_{\|(u, v)\|=1} \|B'(k, f)(u, v) - B'(\tilde{k}, \tilde{f})(u, v)\| \\ &\leq \sqrt{2} C \|(k - \tilde{k}, f - \tilde{f})\|. \end{aligned}$$

□

Note that the adjoint operator $(B'(k, f))^*$ of the Frechét derivative $B'(k, f)$ exists and is a bounded linear operator whenever both \mathcal{H} and $\mathcal{U} \times \mathcal{V}$ are Hilbert spaces.

In order to analyse the optimality condition (9) we shall compute the subdifferential of a functional over two variables. As pointed out in [6, proposition 2.3.15] for a general function h the set-valued mapping $\partial h: \mathcal{U} \rightrightarrows \mathcal{U}^*$ the set $\partial h(x_1, x_2)$ and the product set $\partial_1 h(x_1, x_2) \times \partial_2 h(x_1, x_2)$ are not necessarily contained in each other. Here, $\partial_i h$ denotes the partial subgradient with respect to x_i for $i = 1, 2$. However this is not the case for the functional we are interested in as will be shown in the following theorem.

Theorem 3.2. Let $J: \mathcal{U} \times \mathcal{V} \rightarrow \overline{\mathbb{R}}$ be a functional with the structure

$$J(u, v) = \varphi(u) + Q(u, v) + \psi(v), \quad (10)$$

where Q is a (nonlinear) differentiable term and $\varphi: \mathcal{U} \rightarrow \overline{\mathbb{R}}$, $\psi: \mathcal{V} \rightarrow \overline{\mathbb{R}}$ are proper convex functions, $u \in \text{dom } \varphi$ and $v \in \text{dom } \psi$. Then

$$\begin{aligned} \partial J(u, v) &= \{ \partial \varphi(u) + Q'_u(u, v) \} \times \{ \partial \psi(v) + Q'_v(u, v) \} \\ &= \{ \partial_u J(u, v) \} \times \{ \partial_v J(u, v) \}. \end{aligned}$$

Proof. In general the subdifferential of a sum of functions does not equal the sum of its subdifferentials. However, if Q is differentiable, φ and ψ are convex some inclusions and even equalities hold true (combining [6, proposition 2.3.3; corollary 3; proposition 2.3.6]), as for instance,

$$\partial J(u, v) = \partial(\varphi(u) + \psi(v)) + \partial Q(u, v).$$

Since Q is differentiable, calling the previous results, the (partial) subderivative is unique [6, proposition 2.3.15] and therefore

$$\begin{aligned} \partial Q(u, v) &= \partial_u Q(u, v) \times \partial_v Q(u, v) \\ &= (Q'_u(u, v), Q'_v(u, v)). \end{aligned}$$

Note that the subderivative of the sum of two separable convex functionals satisfies

$$\partial(\varphi(u) + \psi(v)) = (\partial \varphi(u), \partial \psi(v))$$

see [18, corollary 2.4.5].

Altogether, we can compute the subderivative as follows

$$\begin{aligned} \partial J(u, v) &= (\partial \varphi(u), \partial \psi(v)) + (Q'_u(u, v), Q'_v(u, v)) \\ &= \{ \partial_u \varphi(u) + Q'_u(u, v) \} \times \{ \partial_v \psi(v) + Q'_v(u, v) \}. \end{aligned} \quad (11)$$

The last implication of this theorem,

$$\partial J(u, v) = \{ \partial_u J(u, v) \} \times \{ \partial_v J(u, v) \}$$

follows straightforward by the definition of partial subderivative and (11). \square

Please note that the above proof holds for all definitions of subdifferentials introduced in the appendix, as for convex functionals all the definitions are equivalent, and for differentiable (possibly nonlinear) terms the subdifferential is a singleton and the subderivative equals the derivative. Based on theorem 3.2 we can now calculate the derivative of the functional which is the gist for building up the upcoming algorithm; please give heed to the structure of (10) and the proposed functional $J_{\alpha, \beta}^{\delta, \epsilon}$:

Corollary 3.3. Let $J_{\alpha, \beta}^{\delta, \epsilon}$ the functional defined in (8), then

$$\partial J_{\alpha, \beta}^{\delta, \epsilon}(k, f) = \left\{ C_f^*(C_f k - g_\delta) + \gamma(k - k_\epsilon) + \beta \zeta \right\} \times \left\{ A_k^*(A_k f - g_\delta) + \alpha L^* L f \right\}$$

where $\zeta \in \partial \mathcal{R}(k)$.

Proof. The result follows straightforward from lemma 3.1 and theorem 3.2. Observe that the sum $C_f^*(C_f k - g_\delta) + \gamma(k - k_e) + \beta\zeta$ is well-defined in the Hilbert space \mathcal{U} , since the subderivative $\zeta \partial \mathcal{R}(k)$ is also an element of \mathcal{U} . \square

Up to now, we did not specify the functional \mathcal{R} , it is only required to be convex and lower semi-continuous. We are particularly interested in, e.g. the L_p norm or the weighted ℓ_p norm, denoted by $\mathcal{R}(k) = \|k\|_{w,p}$. Its subdifferential is given in section 4. An easy way to compute the subderivatives of functionals \mathcal{R} with a specific structure is given by the following lemma.

Lemma 3.4. [(3, lemma 4.4)] Let $\mathcal{H} = L_2(\Omega, d\mu)$ where Ω is a σ -finite measure space. Let $\mathcal{R}: \mathcal{H} \rightarrow (-\infty, +\infty]$ be defined by

$$\mathcal{R}(u) = \begin{cases} \int_{\Omega} h(u) d\mu & \text{if the integral is finite} \\ \infty & \text{else,} \end{cases} \quad (12)$$

where $h: \mathbb{C} \rightarrow \mathbb{R}$ is a convex function. Then $\xi \in \mathcal{H}$ is an element of $\partial \mathcal{R}(u)$ if and only if $\xi(x) \in \partial h(u(x))$ for almost every $x \in \Omega$ (with the identification $\mathbb{C} = \mathbb{R}^2$).

3.2. An alternating minimisation algorithm

Coordinate descent methods are based on the idea that the minimisation of a multivariable function can be achieved by minimising it along one direction at a time. It is a simple and surprisingly efficient technique. The coordinates can be chosen arbitrarily with any permutation, but one can also replace them by block coordinates (for more details see [16] and references therein). This method is closely related to coordinate gradient descent (CGD), Gauss–Seidel and SOR methods, which was studied previously by several authors and described in various optimisation books, e.g. [1, 14]. In the unconstrained setting the method is called *alternating minimisation* (AM) when the variables are split into two blocks.

The computation of a solution of dbl-RTLS is not straightforward, as determining the minimum of the functional (8) with respect to both parameters is a nonlinear and nonconvex problem over two variables. Nevertheless we shall overcome this problem by applying some coordinate descent techniques.

In the following we shall denote the dbl-RTLS functional by J instead of $J_{\alpha,\beta}^{\delta,\epsilon}$, as the parameters of the functionals are kept fix for the minimisation process.

In the AM algorithm, the functional is minimised iteratively with two alternating minimisation steps. Each step minimises the problem over one variable while keeping the second variable fixed:

$$f^{n+1} \in \arg \min_{f \in V} J(k, f | k^n) \quad (13a)$$

$$k^{n+1} \in \arg \min_{k \in U} J(k, f | f^{n+1}). \quad (13b)$$

The notation $J(k, f | u)$ means we minimise the function J with u fixed, where u can be either k or f . Thus we minimise in each cycle the functionals

$$J(k, f | k^n) = \|A_k f - g_\delta\|^2 + \alpha \|L f\|^2,$$

and

$$J(k, f|f^{n+1}) = \|C_{f^{n+1}}k - g_\delta\|^2 + \gamma \|k - k_\epsilon\|^2 + \beta\mathcal{R}(k).$$

We highlight some important facts:

1. For each subproblem, the considered operators are linear, and the functional is convex. Thus a local minimum is global.
2. The first step is a standard quadratic minimisation problem.

First we will show a monotonicity result for the sequence $\{(k^n, f^n)\}_n$ of iterates:

Proposition 3.5. *The functional J is non-increasing on the AM iterates,*

$$J(k^{n+1}, f^{n+1}) \leq J(k^n, f^{n+1}) \leq J(k^n, f^n).$$

Proof. The iterates are defined as

$$f^{n+1} \in \arg \min_{f \in V} J(k, f|k^n)$$

and

$$k^{n+1} \in \arg \min_{k \in U} J(k, f|f^{n+1}).$$

Therefore,

$$J(k^n, f^{n+1}) \leq J(k^n, f) \quad \forall f \in V$$

and

$$J(k^{n+1}, f^{n+1}) \leq J(k, f^{n+1}) \quad \forall k \in U,$$

and in particular, setting $f = f^n$ and $k = k^n$,

$$\begin{aligned} J(k^n, f^{n+1}) &\leq J(k^n, f^n) \\ J(k^{n+1}, f^{n+1}) &\leq J(k^n, f^{n+1}), \end{aligned}$$

and

$$J(k^{n+1}, f^{n+1}) \leq J(k^n, f^{n+1}) \leq J(k^n, f^n).$$

□

The existence of a minimiser of the functional J has already been proven in [2, theorem 4.2]. The goal of the following results is to prove that the sequence generated by the alternating minimisation algorithm has at least a subsequence which converges towards to a critical point of the functional. Throughout this section, let us make the following assumptions.

Assumption A.

- (A1) B is strongly continuous, i.e. if $(k^n, f^n) \rightarrow (\bar{k}, \bar{f})$ then $B(k^n, f^n) \rightarrow B(\bar{k}, \bar{f})$.
 (A2) The adjoint of the Fréchet derivative B' of B is strongly continuous, i.e. if $(k^n, f^n) \rightarrow (\bar{k}, \bar{f})$ then $B'(k^n, f^n)^*z \rightarrow B'(\bar{k}, \bar{f})^*z, \forall z \in \mathcal{D}(B')$

Additionally to the standard norm for the pair $(k, f) \in \mathcal{U} \times \mathcal{V}$

$$\|(k, f)\|^2 = \|k\|^2 + \|f\|^2$$

we define the weighted norm for given $\gamma > 0$ as

$$\|(k, f)\|_\gamma^2 = \gamma \|k\|^2 + \|f\|^2.$$

Proposition 3.6. For given regularisation parameters $0 < \alpha$ and β , the sequence $\{(k^{n+1}, f^{n+1})\}_{n+1}$ of iterates generated by the AM algorithm has at least a weakly convergent subsequence $(k^{n_j+1}, f^{n_j+1}) \rightarrow (\bar{k}, \bar{f})$, and its limit fulfils

$$J(\bar{k}, \bar{f}) \leq J(\bar{k}, f) \quad \text{and} \quad J(\bar{k}, \bar{f}) \leq J(k, \bar{f}) \quad (14)$$

for all $f \in \mathcal{V}$ and for all $k \in \mathcal{U}$.

Proof. As the iterates of the AM algorithm can be characterised as the minimisers of a reduced dbl-RTLS functional, see (13a), (13b) we observe

$$\begin{aligned} \alpha \|L f^{n+1}\|^2 + \gamma \|k^n - k_\epsilon\|^2 + \beta \mathcal{R}(k^n) &\leq J(k^n, f^{n+1}) \\ &= \min_f J(k, f | k^n) \\ &\leq J(k^n, 0) \\ &= \|g_\delta\|^2 + \gamma \|k^n - k_\epsilon\|^2 + \beta \mathcal{R}(k^n) \end{aligned}$$

and

$$\begin{aligned} \alpha \|L f^{n+1}\|^2 + \gamma \|k^{n+1} - k_\epsilon\|^2 &\leq J(k^{n+1}, f^{n+1}) \\ &= \min_k J(k, f | f^{n+1}) \\ &\leq J(0, f^{n+1}) \\ &= \|g_\delta\|^2 + \gamma \|k_\epsilon\|^2 + \alpha \|L f^{n+1}\|^2. \end{aligned}$$

Keeping in mind that the operator L is continuously invertible, the first inequality gives

$$\|f^{n+1}\|^2 \leq \frac{1}{\|L^{-1}\|^2 \alpha} \|g_\delta\|^2.$$

Using the second estimate above and the standard inequality $\|a + b\|^2 \leq 2(\|a\|^2 + \|b\|^2)$ we have

$$\gamma \left\| k^{n+1} \right\|^2 \leq 2 \left\| g_\delta \right\|^2 + 4\gamma \left\| k_\epsilon \right\|^2.$$

Thus, the sequence $\{(k^{n+1}, f^{n+1})\}_{n+1}$ is bounded

$$\begin{aligned} \left\| (k^{n+1}, f^{n+1}) \right\|_\gamma^2 &= \gamma \left\| k^{n+1} \right\|^2 + \left\| f^{n+1} \right\|^2 \\ &\leq 2 \left\| g_\delta \right\|^2 + 4\gamma \left\| k_\epsilon \right\|^2 + \frac{1}{c^2\alpha} \left\| g_\delta \right\|^2 \\ &= \left(2 + \frac{1}{\left\| L^{-1} \right\|^2 \alpha} \right) \left\| g_\delta \right\|^2 + 4\gamma \left\| k_\epsilon \right\|^2 \end{aligned}$$

and by Alaoglu's theorem, it has a weakly convergent subsequence $\{(k^{n_j+1}, f^{n_j+1})\}_{n_j+1} \rightharpoonup (\bar{k}, \bar{f})$.

Since f^{n_j+1} minimises the functional $J(k^{n_j}, f)$ for a fixed k_j^n , it holds

$$J(k^{n_j}, f^{n_j+1}) \leq J(k^{n_j}, f) \quad \forall f \in \mathcal{V}$$

and thus

$$\left\| B(k^{n_j}, f^{n_j+1}) - g_\delta \right\|^2 + \alpha \left\| Lf^{n_j+1} \right\|^2 \leq \left\| B(k^{n_j}, f) - g_\delta \right\|^2 + \alpha \left\| Lf \right\|^2.$$

Using the fact that J is w-lsc and the strong continuity of B , we observe

$$\begin{aligned} &\left\| B(\bar{k}, \bar{f}) - g_\delta \right\|^2 + \alpha \left\| L\bar{f} \right\|^2 \\ &\leq \liminf_{n_j \rightarrow \infty} \left\{ \left\| B(k^{n_j+1}, f^{n_j+1}) - g_\delta \right\|^2 + \alpha \left\| Lf^{n_j+1} \right\|^2 \right\} \\ &\leq \liminf_{n_j \rightarrow \infty} \left\{ \left\| B(k^{n_j}, f^{n_j+1}) - g_\delta \right\|^2 + \alpha \left\| Lf^{n_j+1} \right\|^2 \right\} \\ &\leq \liminf_{n_j \rightarrow \infty} \left\{ \left\| B(k^{n_j}, f) - g_\delta \right\|^2 + \alpha \left\| Lf \right\|^2 \right\} \\ &\leq \limsup_{n_j \rightarrow \infty} \left\{ \left\| B(k^{n_j}, f) - g_\delta \right\|^2 + \alpha \left\| Lf \right\|^2 \right\} \\ &= \lim_{n_j \rightarrow \infty} \left\{ \left\| B(k^{n_j}, f) - g_\delta \right\|^2 + \alpha \left\| Lf \right\|^2 \right\} \\ &\stackrel{(A1)}{=} \left\| B(\bar{k}, f) - g_\delta \right\|^2 + \alpha \left\| Lf \right\|^2 \end{aligned} \tag{15}$$

Therefore,

$$J(\bar{k}, \bar{f}) \leq J(\bar{k}, f) \quad \forall f \in \mathcal{V}.$$

The second inequality in (14) is proven similarly: since k^{n_j+1} minimises the functional $J(k, f^{n_j+1})$ for fixed f^{n_j+1} it is

$$J(k^{n_j+1}, f^{n_j+1}) \leq J(k, f^{n_j+1}) \quad \forall k \in \mathcal{U},$$

which is equivalent to

$$\begin{aligned} & \left\| B(k^{n_j+1}, f^{n_j+1}) - g_\delta \right\|^2 + \gamma \left\| k^{n_j+1} - k_\epsilon \right\|^2 + \beta \mathcal{R}(k^{n_j+1}) \\ & \leq \left\| B(k, f^{n_j+1}) - g_\delta \right\|^2 + \gamma \left\| k - k_\epsilon \right\|^2 + \beta \mathcal{R}(k). \end{aligned}$$

Again, we observe

$$\begin{aligned} & \left\| B(\bar{k}, \bar{f}) - g_\delta \right\|^2 + \gamma \left\| \bar{k} - k_\epsilon \right\|^2 + \beta \mathcal{R}(\bar{k}) \\ & \leq \liminf_{n_j \rightarrow \infty} \left\{ \left\| B(k^{n_j+1}, f^{n_j+1}) - g_\delta \right\|^2 + \gamma \left\| k^{n_j+1} - k_\epsilon \right\|^2 + \beta \mathcal{R}(k^{n_j+1}) \right\} \\ & \leq \liminf_{n_j \rightarrow \infty} \left\| B(k, f^{n_j+1}) - g_\delta \right\|^2 + \gamma \left\| k - k_\epsilon \right\|^2 + \beta \mathcal{R}(k) \\ & = \lim_{n_j \rightarrow \infty} \left\| B(k, f^{n_j+1}) - g_\delta \right\|^2 + \gamma \left\| k - k_\epsilon \right\|^2 + \beta \mathcal{R}(k) \\ & = \left\| B(k, \bar{f}) - g_\delta \right\|^2 + \gamma \left\| k - k_\epsilon \right\|^2 + \beta \mathcal{R}(k), \end{aligned} \tag{16}$$

and thus

$$J(\bar{k}, \bar{f}) \leq J(k, \bar{f}), \quad \forall k \in \mathcal{U}.$$

□

In summary, the AM algorithm yields a bounded sequence $\{(k^{n+1}, f^{n+1})\}_n$ and hence a weakly convergent subsequence. The next two results extend the convergence on the strong topology, for both $\{k^{n_j+1}\}_{n_j}$ and $\{f^{n_j+1}\}_{n_j}$, respectively.

Proposition 3.7. *Let $\{(k^{n_j+1}, f^{n_j+1})\}_{n_j}$ be a weakly convergent (sub-) sequence generated by the AM algorithm (13), where $k^{n_j+1} \rightharpoonup \bar{k}$ and $f^{n_j+1} \rightharpoonup \bar{f}$. Then there exists a subsequence $\{k^{n_{j_m}+1}\}_{n_{j_m}}$ of $\{k^{n_j+1}\}_{n_j}$ such that $k^{n_{j_m}+1} \rightarrow \bar{k}$ and $0 \in \partial_k J(\bar{k}, \bar{f})$.*

Proof. Inequalities (16) in the proposition 3.6's proof reads

$$\begin{aligned} & \liminf_{n_j \rightarrow \infty} \left\{ \left\| B(k^{n_j+1}, f^{n_j+1}) - g_\delta \right\|^2 + \gamma \left\| k^{n_j+1} - k_\epsilon \right\|^2 + \beta \mathcal{R}(k^{n_j+1}) \right\} \\ & = \left\| B(k, \bar{f}) - g_\delta \right\|^2 + \gamma \left\| k - k_\epsilon \right\|^2 + \beta \mathcal{R}(k). \end{aligned}$$

for any k . Setting $k = \bar{k}$ yields in particular

$$\begin{aligned} & \liminf_{n_j \rightarrow \infty} \left\{ \left\| B(k^{n_j+1}, f^{n_j+1}) - g_\delta \right\|^2 + \gamma \left\| k^{n_j+1} - k_\epsilon \right\|^2 + \beta \mathcal{R}(k^{n_j+1}) \right\} \\ & = \left\| B(\bar{k}, \bar{f}) - g_\delta \right\|^2 + \gamma \left\| \bar{k} - k_\epsilon \right\|^2 + \beta \mathcal{R}(\bar{k}). \end{aligned}$$

As the limes inferior exists, we can in particular extract a subsequence $(k^{n_{j_m}+1}, f^{n_{j_m}+1})_{n_{j_m}}$ of $(k^{n_j+1}, f^{n_j+1})_{n_j}$ such that

$$\begin{aligned} & \lim_{n_{j_m} \rightarrow \infty} \left\{ \left\| B(k^{n_{j_m}+1}, f^{n_{j_m}+1}) - g_\delta \right\|^2 + \gamma \left\| k^{n_{j_m}+1} - k_\epsilon \right\|^2 + \beta \mathcal{R}(k^{n_{j_m}+1}) \right\} \\ &= \left\| B(\bar{k}, \bar{f}) - g_\delta \right\|^2 + \gamma \left\| \bar{k} - k_\epsilon \right\|^2 + \beta \mathcal{R}(\bar{k}). \end{aligned} \quad (17)$$

For the sake of notation simplicity we denote for the remainder of the proof the index $n_{j_m} + 1$ by $m + 1$. By (A1) we observe

$$\lim_{m \rightarrow \infty} \left\| B(k^{m+1}, f^{m+1}) - g_\delta \right\|^2 \stackrel{(A1)}{=} \left\| B(\bar{k}, \bar{f}) - g_\delta \right\|^2$$

As all summands in (17) are positive, we have thus and

$$\begin{aligned} \lim_{m \rightarrow \infty} \left\{ \gamma \left\| k^{m+1} - k_\epsilon \right\|^2 + \beta \mathcal{R}(k^{m+1}) \right\} &= \gamma \lim_{m \rightarrow \infty} \left\| k^{m+1} - k_\epsilon \right\|^2 + \beta \lim_{m \rightarrow \infty} \mathcal{R}(k^{m+1}) \\ &= \gamma \left\| \bar{k} - k_\epsilon \right\|^2 + \beta \mathcal{R}(\bar{k}). \end{aligned} \quad (18)$$

Now let us show that k^{m+1} converges strongly. As the sequence converges weakly, it is enough to show

$$\lim_{m \rightarrow \infty} \left\| k^{m+1} \right\|^2 = \left\| \bar{k} \right\|^2$$

Equivalently, we can also show $\lim_{m \rightarrow \infty} \left\| k^{m+1} - k_\epsilon \right\|^2 = \left\| \bar{k} - k_\epsilon \right\|^2$. Again due to the weak convergence of k^{m+1} it is sufficient to prove

$$\limsup_{m \rightarrow \infty} \left\| k^{m+1} - k_\epsilon \right\|^2 \leq \left\| \bar{k} - k_\epsilon \right\|^2.$$

Let us assume that

$$\mu := \limsup_{m \rightarrow \infty} \left\| k^{m+1} - k_\epsilon \right\|^2 > \left\| \bar{k} - k_\epsilon \right\|^2.$$

holds. Rewriting (18) yields

$$\begin{aligned} \beta \limsup_{m \rightarrow \infty} \left\{ \mathcal{R}(k^{m+1}) \right\} &= \gamma \left(\left\| \bar{k} - k_\epsilon \right\|^2 - \limsup_{m \rightarrow \infty} \left\| k^{m+1} - k_\epsilon \right\|^2 \right) + \beta \mathcal{R}(\bar{k}) \\ &= \gamma \left(\left\| \bar{k} - k_\epsilon \right\|^2 - \mu \right) + \beta \mathcal{R}(\bar{k}) \\ &< \beta \mathcal{R}(\bar{k}). \end{aligned} \quad (19)$$

However, since \mathcal{R} is w-lsc, we observe

$$\mathcal{R}(\bar{k}) \leq \liminf_{m \rightarrow \infty} \mathcal{R}(k^{m+1}) \leq \limsup_{m \rightarrow \infty} \mathcal{R}(k^{m+1}),$$

which is in contradiction to (19). Thus we have shown the convergence of k^{m+1} to \bar{k} in norm.

The last part of this proof focus on the convergence of the partial subdifferential of J with respect to k .

Since k^{m+1} solves the sub-minimisation problem (13b), the optimality condition reads as $0 \in \partial_k J(k^{m+1}, f^{m+1})$, or equivalently, there exists an element

$$\xi_k^{m+1} := -\frac{1}{\beta} \left(C_{f^{m+1}}^* (C_{f^{m+1}} k^{m+1} - g_\delta) + \gamma (k^{m+1} - k_\epsilon) \right) \quad (20)$$

such that $\xi_k^{m+1} \in \partial \mathcal{R}(k^{m+1}) \subset \mathcal{U}$; see corollary 3.3.

Now, on the limit, $0 \in \partial_k J(\bar{k}, \bar{f})$, means that

$$\bar{\xi} := -\frac{1}{\beta} \left(C_{\bar{f}}^* (C_{\bar{f}} \bar{k} - g_\delta) + \gamma (\bar{k} - k_\epsilon) \right) \quad \text{and} \quad \bar{\xi} \in \partial \mathcal{R}(\bar{k})$$

holds, i.e. the right hand-side of (20) converges and the limit of the sequence of subderivatives belongs also to the subdifferential set $\partial \mathcal{R}(\bar{k})$.

The first part of the statement above can be seen by using condition (A2). Whereas the second part is obtained by the assumption that \mathcal{R} is a convex functional, because in this case the Fenchel subdifferential coincides with the limiting subdifferential, which is a strong-weakly closed mapping (see appendix).

The strong convergence for the second variable is obtained as follows.

Proposition 3.8. *Let $\{m\}$ be a subsequence of \mathbb{N} such that the (sub-) sequence $\{(k^{m+1}, f^{m+1})\}_m$ generated by AM algorithm (13) satisfies $k^{m+1} \rightarrow \bar{k}$ and $f^{m+1} \rightharpoonup \bar{f}$. Then there is a subsequence of $\{f^{m+1}\}_m$ such that $f^{m_j+1} \rightarrow \bar{f}$ and $0 \in \partial_f J(\bar{k}, \bar{f})$.*

Proof. Similarly as the previous theorem, by setting $f = \bar{f}$ at (15) in the proposition 3.6's proof we obtain

$$\begin{aligned} & \liminf_{m \rightarrow \infty} \left\{ \left\| B(k^{m+1}, f^{m+1}) - g_\delta \right\|^2 + \alpha \left\| Lf^{m+1} \right\|^2 \right\} \\ &= \left\| B(\bar{k}, \bar{f}) - g_\delta \right\|^2 + \alpha \left\| L\bar{f} \right\|^2. \end{aligned}$$

As the limes inferior exists, we can in particular extract a subsequence $(k^{m_j+1}, f^{m_j+1})_{m_j}$ of $(k^{m+1}, f^{m+1})_m$ such that

$$\begin{aligned} & \lim_{m_j \rightarrow \infty} \left\{ \left\| B(k^{m_j+1}, f^{m_j+1}) - g_\delta \right\|^2 + \alpha \left\| Lf^{m_j+1} \right\|^2 \right\} \\ &= \left\| B(\bar{k}, \bar{f}) - g_\delta \right\|^2 + \alpha \left\| L\bar{f} \right\|^2. \end{aligned}$$

Since both summands in the limit above are positive and due to (A1), we conclude that

$$\lim_{m_j \rightarrow \infty} \left\| Lf^{m_j+1} \right\|^2 = \left\| L\bar{f} \right\|^2.$$

Moreover, as L is a bounded and continuously invertible operator we have

$$\lim_{m_j \rightarrow \infty} \left\| f^{m_j+1} \right\|^2 = \left\| \bar{f} \right\|^2,$$

which in combination with the weak convergence of the subsequence gives its strong convergence $f^{m_j+1} \rightarrow \bar{f}$.

The second half of this proof refers to the convergence of the partial subdifferential of J with respect to f and its limit.

Since f^{m+1} solves the sub-minimisation problem (13a), the optimality condition reads as $0 \in \partial_f J(k^m, f^{m+1})$. However we are interested on the partial subderivate at the pair

(k^{m_j+1}, f^{m_j+1}) . Namely, with help of corollary 3.3 the subderivative (which is a unique element) $\xi_f^{m_j+1} \in \partial_f J(k^{m_j+1}, f^{m_j+1})$ is computed³ as

$$\xi_f^{m+1} := A_{k^{m+1}}^* (A_{k^{m+1}} f^{m+1} - g_\delta) + \alpha L^* L f^{m+1},$$

which may not be necessarily null for each cycle of the AM algorithm (13), otherwise the stopping criteria would be satisfied and nothing would be left to be proven. Therefore we shall prove that it converges towards zero.

So far we have strong convergence of both sequences $\{k^{m+1}\}_m$ and $\{f^{m+1}\}_m$. Additionally, the assumption A implies that both linear operators A_k and A_k^* are also strongly continuous, therefore

$$\begin{aligned} \lim_{m \rightarrow \infty} \xi_f^{m+1} &= \lim_{m \rightarrow \infty} \left\{ A_{k^{m+1}}^* (A_{k^{m+1}} f^{m+1} - g_\delta) + \alpha L^* L f^{m+1} \right\} \\ &= A_{\bar{k}}^* (A_{\bar{k}} \bar{f} - g_\delta) + \alpha L^* L \bar{f}. \end{aligned} \quad (21)$$

Our goal is to show that the limit given in (21) is zero. Let's suppose by contradiction that $0 \notin \partial_f J(\bar{k}, \bar{f})$. Since this set is singleton we conclude that

$$A_{\bar{k}}^* (A_{\bar{k}} \bar{f} - g_\delta) + \alpha L^* L \bar{f} \neq 0.$$

This means that \bar{f} does not fulfil the normal equation associated to the standard Tikhonov problem

$$\text{minimise}_f \left\| A_{\bar{k}} f - g_\delta \right\|^2 + \alpha \|L f\|^2,$$

which is a necessary condition to be a minimiser candidate to the underlying functional.

Therefore the functional $J(\bar{k}, \cdot)$ for a given fixed \bar{k} does not attain its minimum value at \bar{f} and there is at least one element f such that $J(\bar{k}, f) < J(\bar{k}, \bar{f})$.

Moreover this functional is convex and it has a global solution, here denoted by \tilde{f} . By definition

$$J(\bar{k}, \tilde{f}) \leq J(\bar{k}, f)$$

for all $f \in V$.

In particular, since \bar{f} is not a minimiser for $J(\bar{k}, \cdot)$, the inequality above is strict,

$$J(\bar{k}, \tilde{f}) < J(\bar{k}, \bar{f}). \quad (22)$$

On the other hand, from proposition 3.6 it also holds

$$J(\bar{k}, \bar{f}) \leq J(\bar{k}, f)$$

for all $f \in V$. Setting $f := \tilde{f}$ in this inequality we get

$$J(\bar{k}, \bar{f}) \leq J(\bar{k}, \tilde{f}),$$

which leads to a contradiction to (22).

Therefore for \bar{f} the optimality condition holds true, i.e. in the limit the source condition is fulfilled and the limit of the partial subderivative sequence is zero, i.e. $0 \in \partial_f J(\bar{k}, \bar{f})$, which completes the proof. \square

³ For sake of notation we continue to denote the subsequence's indices by $m + 1$ instead of $m_j + 1$.

Remark 3.9. One alternative proof would be assuming that the sequence $\{k^{m+1}\}_m$ fulfils

$$\|k^{m+1} - k^m\| \rightarrow 0. \quad (23)$$

More specifically, we have

$$A_{k^m}^* (A_{k^m} f^{m+1} - g_\delta) + \alpha L^* L f^{m+1} = 0$$

from the optimality condition, but we would like to show

$$\lim_{m \rightarrow \infty} \left\{ A_{k^{m+1}}^* (A_{k^{m+1}} f^{m+1} - g_\delta) + \alpha L^* L f^{m+1} \right\} = 0.$$

Subtracting the latter expression from the first one, we get

$$(A_{k^m}^* A_{k^m} - A_{k^{m+1}}^* A_{k^{m+1}}) f^{m+1} + (A_{k^m}^* - A_{k^{m+1}}^*) g_\delta.$$

Note that by assuming the condition (23) the expression above converges to zero and the proof would be complete. Nevertheless we cannot guarantee that subsequent elements of the original sequence will be selected for the subsequence. As an alternative one can verify numerically if the sequence provided from the AM algorithm satisfies this assumption. Moreover, if we restrict the problem to the simple case that the characterising function is known, then the assumption (23) is trivial, the problem becomes the standard Tikhonov regularisation and the theory is carried on.

The forthcoming and most substantial result within this section shows that the limit (\bar{k}, \bar{f}) of the sequence generated by the AM algorithm is a critical point (pair) of the functional J .

Theorem 3.10 (Main result). *Let $\{m\}$ an index set of \mathbb{N} such that the sequence generated by AM algorithm $\{(k^{m+1}, f^{m+1})\}_m \rightarrow (\bar{k}, \bar{f})$ and $(\xi_k^{m+1}, \xi_f^{m+1}) \rightarrow (0, 0)$. Then there is subsequence converging towards to a critical point of J , i.e.*

$$(0, 0) \in \partial J(\bar{k}, \bar{f}).$$

Proof. The proposition 3.7 guarantees that $k^{m+1} \rightarrow \bar{k}$ and $\xi_{k^{m+1}} \in \partial \mathcal{R}(k^{m+1})$ (or equivalently, $0 \in \partial_k J(k^{m+1}, f^{m+1})$) such that $0 \in \partial_k J(\bar{k}, \bar{f})$. Likewise, proposition 3.8 guarantees that the sequence $f^{m+1} \rightarrow \bar{f}$ and $\xi_{f^{m+1}} \in \partial J(k^{m+1}, f^{m+1})$ such that $0 \in \partial_f J(\bar{k}, \bar{f})$. Combining this with the strong-weakly closedness property of the subderivative (see appendix) and theorem 3.2 we have

$$(0, 0) \in \partial J(\bar{k}, \bar{f}) = \partial_k J(\bar{k}, \bar{f}) \times \partial_f J(\bar{k}, \bar{f})$$

on the limit. □

4. Numerical experiments

In the previous sections we proposed an algorithm to minimise the functional J over two variables. Here we want to discuss the practical realisation of the algorithm, which has been implemented in MATLAB.

For our test computations we choose \mathcal{R} to be the weighted l_p norm of the coefficients of the characterising function k with respect to an orthonormal basis $\{\phi_\lambda\}_\lambda$ of \mathcal{U} , so

$$\|k\|_{w,p}^p := \sum_\lambda w_\lambda |k_\lambda|^p, \quad (24)$$

where $k_\lambda = |\langle k, \phi_\lambda \rangle|$. For all possible choices of p it is well known that the choice $p = 1$ promotes sparsity [8], in the sense that the minimiser of the related Tikhonov functional has only few nonzero coefficients with respect to the underlying bases. We are particularly interested in wavelets bases, as many signals (1D) or images (2D) exhibit piecewise smooth behaviour punctuated by transients.

One cycle of the alternating minimisation problem (13) consists of two steps, where each step consists of the minimisation of a linear and convex functional. Firstly, solving (13a) we fix k^n and find the solution f^{n+1} through, e.g. a conjugate gradient method. Secondly, solving (13b) we fix f^{n+1} from the previous step and solve the *shrinkage minimisation* problem described on [8] and get k^{n+1} .

We test the performance of our algorithm for a 2D convolution problem

$$f^{\text{true}} * k_0 = g_0.$$

More precisely the image f^{true} is represented numerically as a matrix of size 256×256 , using the command `imread` to read the original JPEG image⁴ file composed by three levels of grey. The blurring kernel k_0 is described by a Gaussian function

$$k_0(x, y) = a \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2c^2}\right)$$

evaluated in the mesh $[-\frac{\pi}{2}, \frac{\pi}{2}] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$, where a is the amplitude constant, $c \in \mathbb{R}$ and x_0 and y_0 define the centre, which in our case is located in the upper-left corner. Finally the convolution operator is evaluated using the fast Fourier transform in 2D and it creates a shifted and blurred image as seen in the figure 1.

Numerical experiments are performed from given measurements in order to reconstruct both the function and the kernel. An example of the initial noisy data and noisy kernel is illustrated on figure 2, where we added 8% relative white noise.

The figure 3 illustrates the significant improvement from the initial given noisy data (see figure 2 (right) with 8% relative noise, $\text{SNR} = 4.199$) compared to the one obtained from the dbI-RTLS solution. For comparison, we also give a reconstruction result obtained by using Tikhonov regularisation applied to the linear convolution problem where the noisy kernel is fixed. The reconstruction result shows that our approach which considers both the function and the kernel as variable leads to a better reconstruction. However, this effect becomes less prominent when the noise becomes considerably small.

The numerical results are given in figure 4, which displays in each row three graphics: the approximated image, the reconstructed kernel and its convolution. It plots a collection of numerical solutions computed from four samples with 8%, 4%, 2% and 1% relative error (RE) on both measurements, respectively in each row from top to bottom. Moreover, we compare the numerical reconstruction with the true image and kernel; the errors in norm are displayed in the table 1. Either numerically or visually one can conclude that dbI-RTLS is indeed a regularisation method, since its reconstruction and computed data improve as the noise level decreases.

⁴ DK Computational Mathematics' logo from JKU Linz.

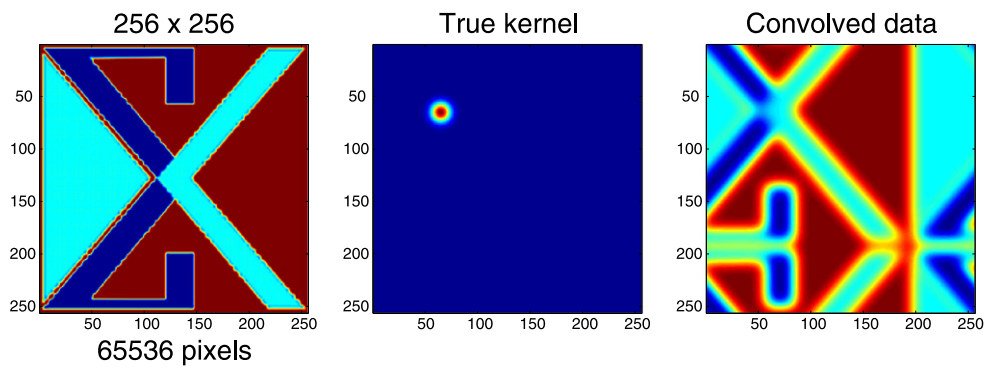


Figure 1. From left to right: true image f^{true} , blurring Gaussian kernel k_0 and convolved data g_0 .

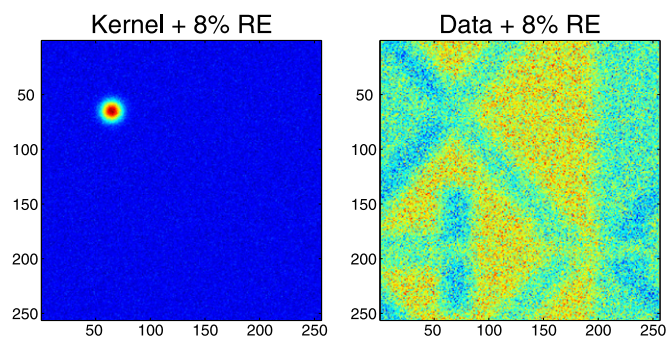


Figure 2. Measurements: noisy kernel (left) and noisy data (right), both with 8% relative white noise error.

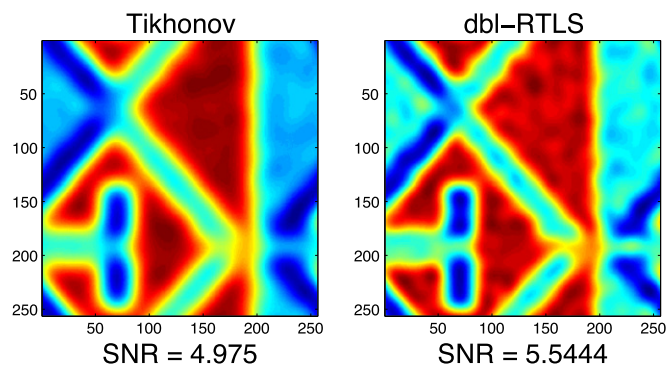


Figure 3. Reconstruction with standard Tikhonov using the blurred kernel (left) and reconstruction from the dbl-RTLS method (right).

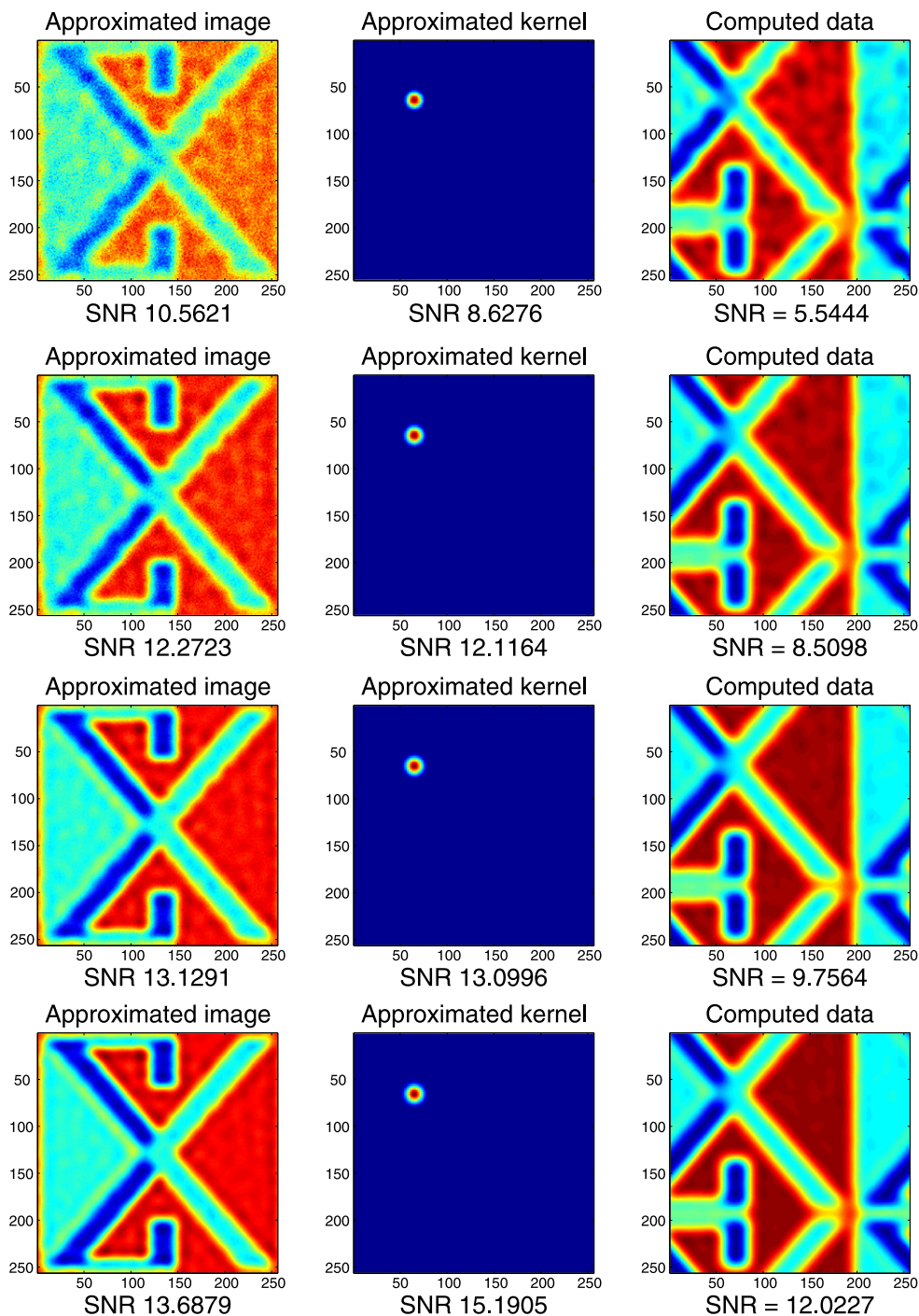


Figure 4. From left to right columns: deconvolution solution f^n , the reconstruction of the characterising function k^n and the attained data g^n . From the top to bottom each row is the solution given by the AM algorithm initiated with 8%, 4%, 2% and 1% relative error for both g_δ and k_e .

Table 1. Error with 2-norm and respective SNR (signal-to-noise ratio).

RE (k_ϵ)	RE (g_δ)	$\ k_n - \bar{k}\ _2$	$\ f_n - \bar{f}\ _2$	SNR f_n	SNR k_n	β	α
8%	8%	3.6438e-01	1.7311e-01	8.6276	10.562	0.4525	0.1246
4%	4%	2.4185e-01	1.5036e-01	12.116	12.272	0.2262	0.0784
2%	2%	2.1545e-01	1.3648e-01	13.099	13.129	0.1131	0.0493
1%	1%	1.6754e-01	1.2596e-01	15.190	13.687	0.0565	0.0310

The pair of regularisation parameters (α, β) was chosen as

$$\alpha = c_1 \sqrt{2 \log(MN)} (\delta + \epsilon)^\mu$$

$$\beta = c_2 \sqrt{2 \log(MN)} (\delta + \epsilon)^\nu,$$

where $c_1, c_2, 0 < \mu, \nu \leq 1$ have been picked heuristically. The $M \times N$ matrix represents the underlying function; in our numerical example $M = N = 256$.

Depending on the noise level, up to 5 AM cycles has been carried out. Each of the cycles has been stopped whenever the related norms of the computed updates were below the threshold $1e-4$.

Acknowledgments

The authors would like to thank Dr E Resmerita for providing valuable references and helpful comments. The research was funded by the Austrian Science Fund (FWF): W1214-N15, project DK8.

Appendix

The most common concept of subderivative is addressed to convex functions. It was introduced by Fenchel, Moreau and Rockafellar in early 1960 s, but it became popular after [15]. The *Fenchel subdifferential* of a convex function $\varphi: \mathcal{U} \rightarrow \overline{\mathbb{R}}$ (or $[-\infty, +\infty]$) at $\bar{u} \in \mathcal{U}$ is defined as the set

$$\partial^F \varphi(\bar{u}) = \left\{ \xi \in \mathcal{U}^* \mid \varphi(\bar{u} + d) - \varphi(\bar{u}) \geq \langle \xi, d \rangle \quad \forall d \in \mathcal{U} \right\}.$$

This definition was extended also to nonconvex functions by Clarke in 1973. It is based on generalised directional derivatives for locally Lipschitz functions in Banach spaces [6]. The *Clark subdifferential* of φ at \bar{u} is defined by

$$\partial^C \varphi(\bar{u}) = \left\{ \xi \in \mathcal{U}^* \mid \varphi^\circ(\bar{u}; d) \geq \langle \xi, d \rangle \quad \forall d \in \mathcal{U} \right\}$$

where

$$\varphi^\circ(\bar{u}; d) = \limsup_{\substack{u \rightarrow \bar{u} \\ t \downarrow 0}} \frac{\varphi(u + td) - \varphi(u)}{t}$$

is the *generalised directional derivative*.

We add to this list two more definitions of subdifferentials. As before, for a set-valued mapping $G: \mathcal{U} \rightrightarrows \mathcal{U}^*$ between a Banach space \mathcal{U} and its topological dual \mathcal{U}^* , the set

$$\text{Lim sup}_{u \rightarrow \bar{u}} G(\bar{u}) = \left\{ \xi \in \mathcal{U}^* \mid \exists u^n \rightarrow \bar{u} \text{ and } \xi^n \xrightarrow{*} \xi \text{ with } \xi^n \in G(u^n) \quad \forall n \in \mathbb{N} \right\}$$

denotes the *sequential Painlevé-Kuratowski upper/outer limit* of a set-valued mapping. Given a lower semi-continuous function φ , the ε -Fréchet subdifferential of φ at \bar{u} is defined by

$$\hat{\partial}_\varepsilon \varphi(\bar{u}) = \left\{ \xi \in \mathcal{U}^* \mid \liminf_{\|d\| \rightarrow 0} \frac{\varphi(\bar{u} + d) - \varphi(\bar{u}) - \langle \xi, d \rangle}{\|d\|} \geq \varepsilon \right\}.$$

If $|\varphi(\bar{u})| = \infty$ then $\hat{\partial}_\varepsilon \varphi(\bar{u}) = \emptyset$. When $\varepsilon = 0$ the set $\hat{\partial}_0 \varphi(\bar{u})$ will be denoted by $\hat{\partial} \varphi(\bar{u})$.

The *limiting subdifferential* or *Mordukhovich subdifferential* of φ at \bar{u} is defined as

$$\partial \varphi(\bar{u}) = \text{Lim sup}_{\substack{u \xrightarrow{\varphi} \bar{u} \\ \varepsilon \downarrow 0}} \hat{\partial}_\varepsilon \varphi(\bar{u})$$

where the notation $u \xrightarrow{\varphi} \bar{u}$ means $u \rightarrow \bar{u}$ with $\varphi(u) \rightarrow \varphi(\bar{u})$. This subdifferential corresponds to the collection of weak-star sequential limiting points of the so-called ε -Fréchet subdifferential.

In [7], the following inclusion property between the sets

$$\partial^F \varphi(\bar{u}) \subset \hat{\partial} \varphi(\bar{u}) \subset \partial^C \varphi(\bar{u}).$$

is shown. The set of subgradients $\hat{\partial} \varphi(\bar{u})$ may be nonconvex, whereas the Clark subdifferential is always a nonempty convex subset of \mathcal{U}^* whenever $\bar{u} \in \text{dom } \varphi$. It is important to note that *the subdifferential definitions generate the same set if the function is convex* [5].

Finally we list another property needed to prove convergence results: the concept of strong-weak* *closeness* (also called *sw*-closed*) property of the subdifferential mapping's graph.

Given the subdifferential $\partial \varphi$ of a proper lower semi-continuous function φ , saying its graph is *sw*-closed* means whenever $(u^n, \zeta^n) \in \text{Gph } \partial \varphi$ converges in the *sw*-topology* to $(\bar{u}, \bar{\zeta})$ it implies $(\bar{u}, \bar{\zeta}) \in \text{Gph } \partial \varphi$. In other words, if $u^n \rightarrow \bar{u}$ and $\zeta^n \xrightarrow{*} \bar{\zeta}$ with $\zeta^n \in \partial \varphi(u^n)$ then $\bar{\zeta} \in \partial \varphi(\bar{u})$.

The subdifferential is indeed a *sw*-closed* set-value mapping, see for instance [6, proposition 2.1.5] or [9, corollary 5.1]. Moreover, this result holds true for any *maximal monotone point-to-set mapping* and not only for the subdifferential set-value mapping case; see [4, chapter 4].

For more details on the different types of subdifferential and its properties we refer to [6, 9, 13, 15] and references therein.

References

- [1] Bertsekas D P 1999 *Nonlinear Programming* (Nashua, NH: Athena Scientific)
- [2] Bleyer I R and Ramlau R 2013 A double regularization approach for inverse problems with noisy data and inexact operator *Inverse Problems* **29** 16
- [3] Bredies K, Lorenz D A and Maass P 2005 Mathematical concepts of multiscale smoothing *Appl. Comput. Harmon. Anal.* **19** 141–61
- [4] Burachik R S and Iusem A N 2008 *Set-Valued Mappings and Enlargements of Monotone Operators (Springer Optimization and its Applications)* (New York: Springer)
- [5] Chieu N H 2009 The Fréchet and limiting subdifferentials of integral functionals on the spaces $L_1(\Omega, E)$ *J. Math. Anal. Appl.* **360** 704–10

- [6] Clarke F H 1990 *Optimization and Nonsmooth Analysis (Classics in Applied Mathematics vol 5)* 2nd edn (Philadelphia: SIAM)
- [7] Correa R, Jofré A and Thibault L 1992 Characterization of lower semicontinuous convex functions *Proc. Amer. Math. Soc.* **116** 67–72
- [8] Daubechies I, Defrise M and Mol C D 2004 An iterative thresholding algorithm for linear inverse problems with a sparsity constraint *Comm. Pure Appl. Math.* **57** 1413–57
- [9] Ekeland I and Témam R 1999 *Convex Analysis and Variational Problems (Classics in Applied Mathematics vol 28)* (Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM)) Translated from the French
- [10] Golub G H, Hansen P C and O’leary D P 1999 Tikhonov regularization and total least squares *SIAM J. Matrix Anal. Appl.* **21** 185–94
- [11] Lampe J and Voss H 2010 Solving regularized total least squares problems based on eigenproblems *Taiwanese J. Math.* **14** 885–909
- [12] Lu S, Pereverzev S V and Tautenhahn U 2008 Dual Regularized Total Least Squeres and Multi-Parameter Regularization *Computational methods in applied mathematics* **8** 253–62
- [13] Mordukhovich B S and Shao Y H 1995 On nonconvex subdifferential calculus in Banach spaces *J. Convex Anal.* **2** 211–27
- [14] Ortega J M and Rheinboldt W C 2000 *Iterative Solution of Nonlinear Equations in Several Variables (Classics in Applied Mathematics)* (Philadelphia, PA: Society for Industrial and Applied Mathematics)
- [15] Rockafellar R T 1966 Characterization of the subdifferentials of convex functions *Pacific J. Math* **17** 497–510
- [16] Tseng P 2001 Convergence of a block coordinate descent method for nondifferentiable minimization *J. Optim. Theory Appl.* **109** 475–94
- [17] Yu-Li You and Kaveh M 1996 A regularization approach to joint blur identification and image restoration *IEEE Trans. Image Process.* **5** 416–28
- [18] Zălinescu C 2002 *Convex Analysis in General Vector Spaces* (World Scientific)